XILINX
ALL PROGRAMMABLE™

UltraSCALE™
Architecture

High Capacity and High Performance 20nm FPGAs

Steve Young, Dinesh Gaitonde
August 2014

Not a Complete Product Overview

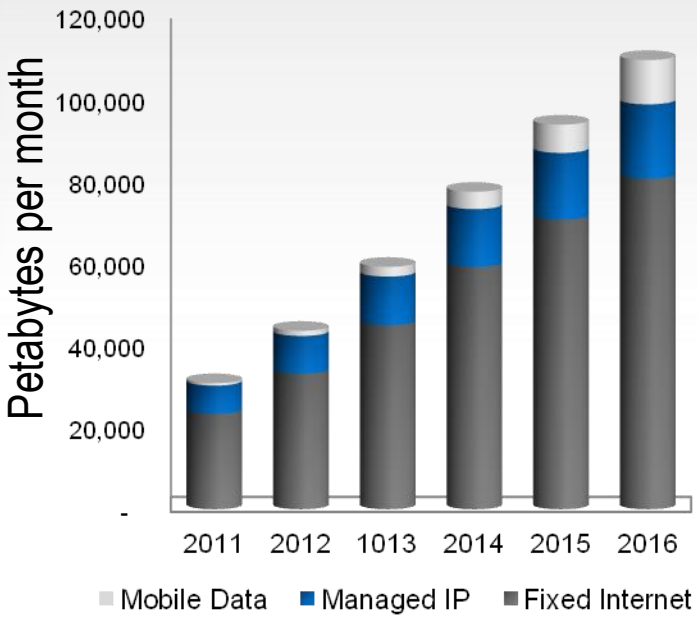
- Doesn’t cover the many new features of UltraScale
- Presentation focused on bandwidth driven fabric challenges at 20nm
- See Xilinx literature for more on UltraScale

Outline

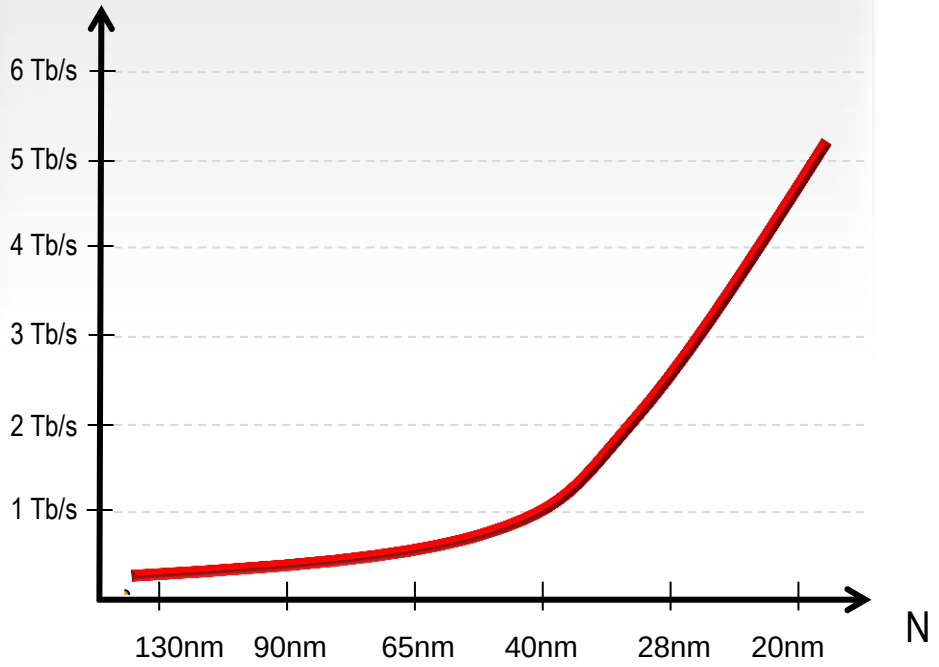
- Motivation: Increasing bandwidth need
- Clocking
- Routing
- Logic density
- Stacked silicon
- Power
- Takeaways

Increasing Bandwidth

Global IP Traffic Growth 2011–2016[1]



Growth of Xilinx FPGA Serial Bandwidth



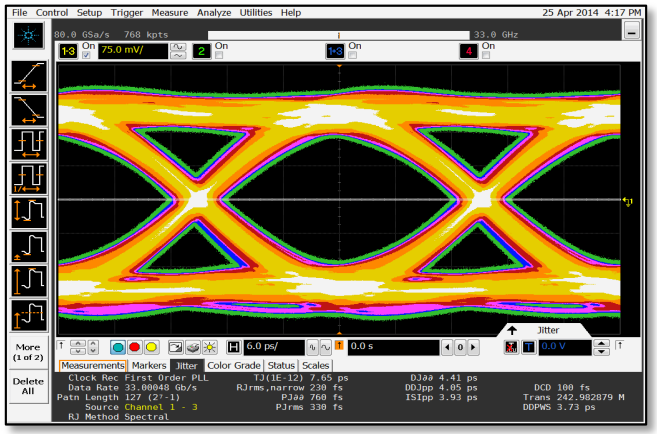
1: Cisco Visual Networking Index, Forecast and Methodology, 2011–2016

Higher Bandwidth IO

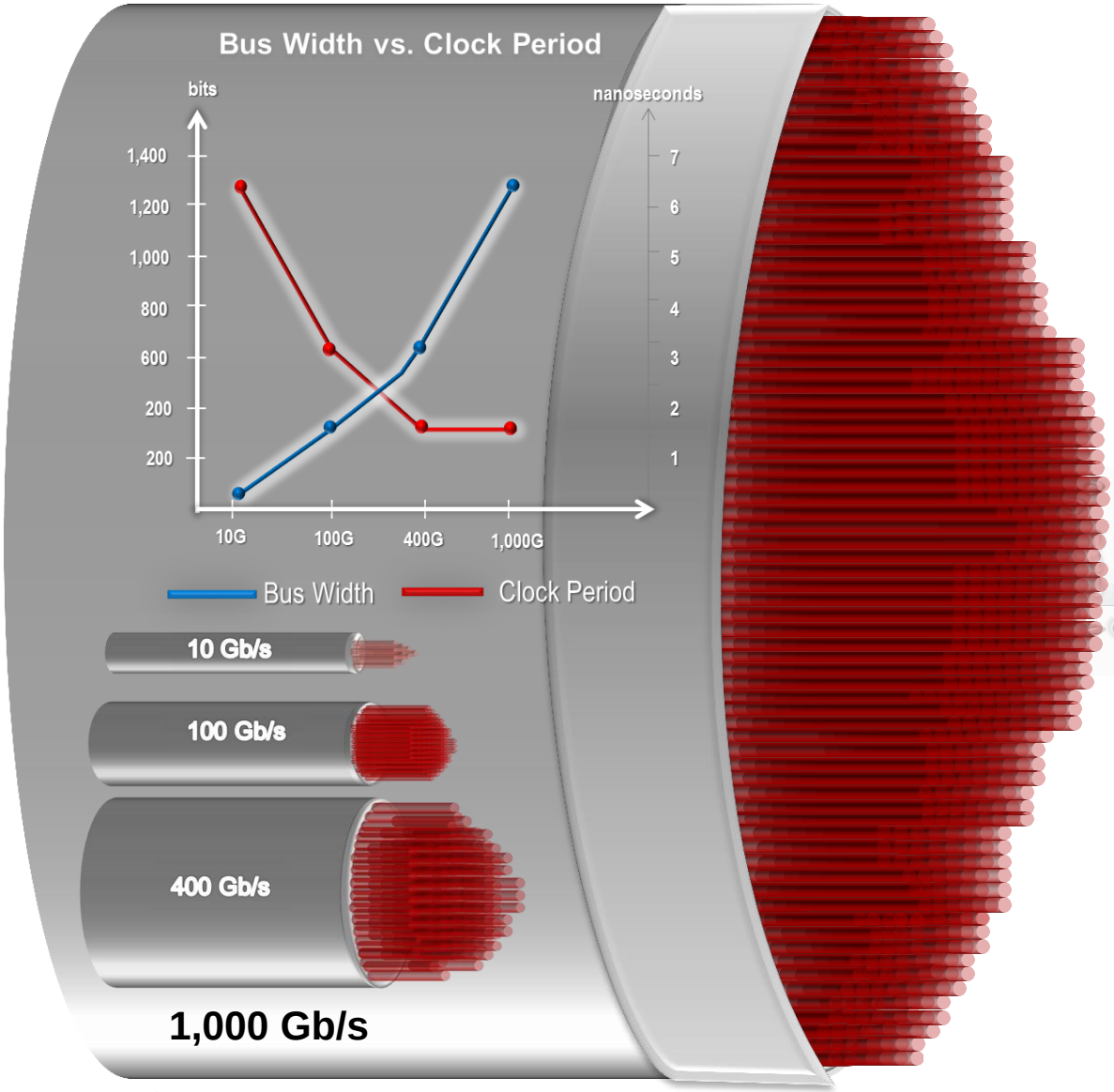
1. Faster transceivers to keep up with increased bandwidth demand, support latest standards
   - Up to 32.75 Gb/s

2. Greater percentage of pins for high speed transciever vs. conventional IO
   - Up to 120 serial transceivers in a device

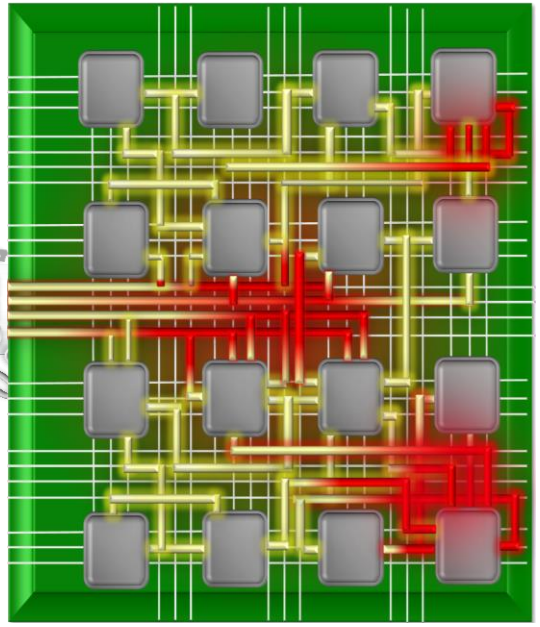
3. Support for up to 6 PCIe Gen3 x8



High Throughput Demands Wider, Faster Data Paths



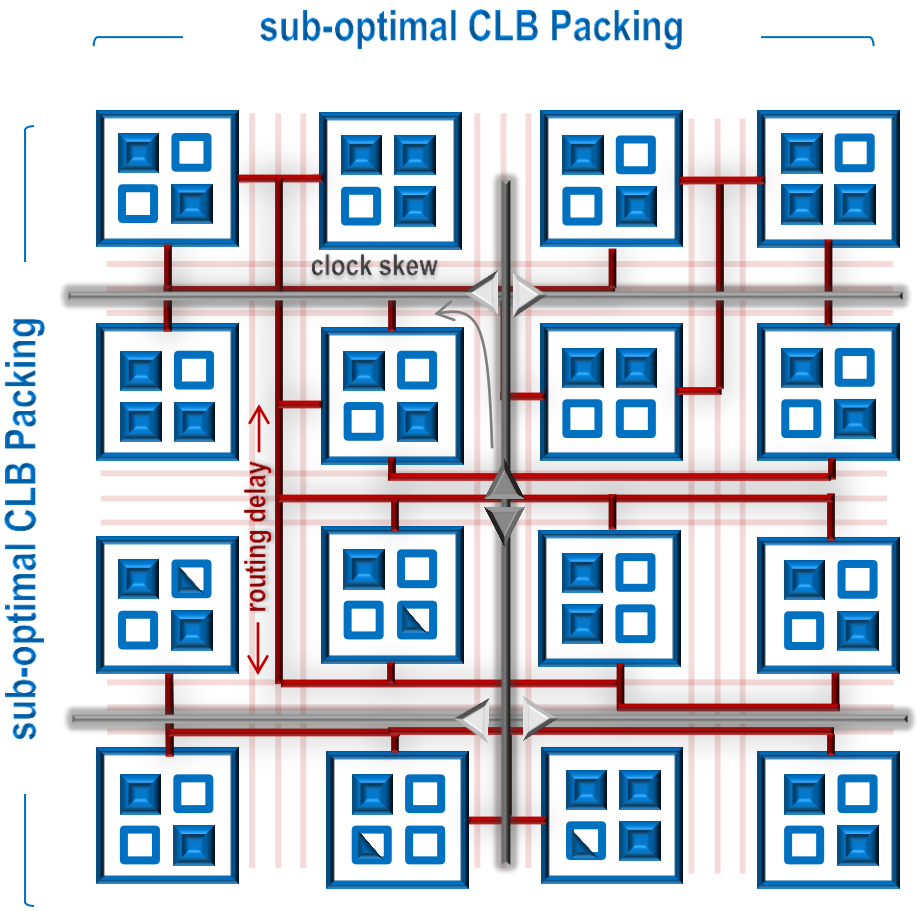
Parallelization of Data Flow



Creates Strain on FPGA Interconnect

Addressing Interconnect Bottlenecks

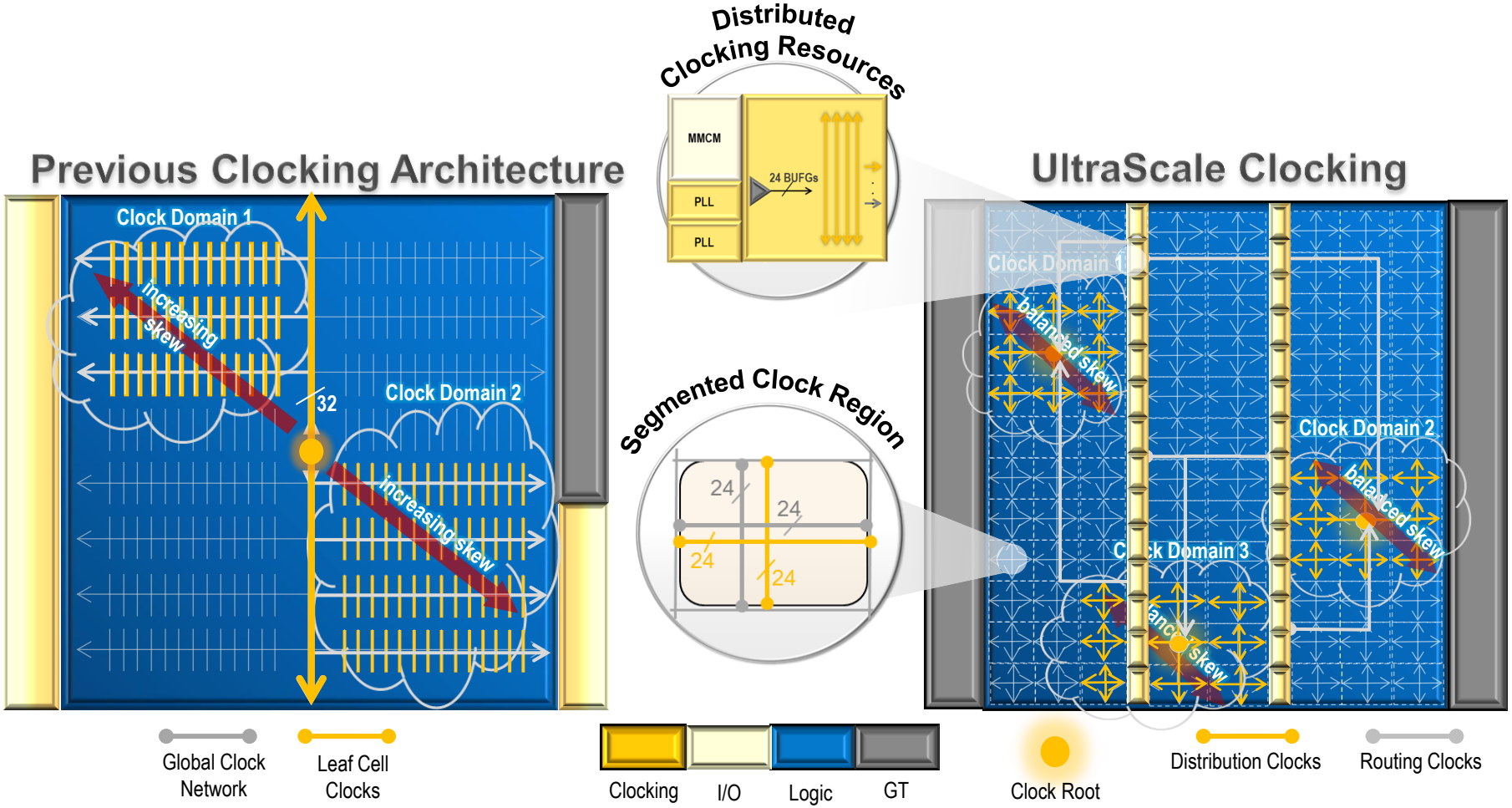
1. **Routing delay** important to overall delay
2. **Clock skew** consumes more timing margin
3. Sub-optimal **CLB packing** reduces performance and utilization



Clocking Challenge

1. Skew
   - As cycle times decrease, Clock skew consumes a bigger percentage of timing margin
2. Quantity
   - As designs grow in size and modularity the number of unique clocks grows and become less global
3. Resistance
   - As metal tracks shrink, metal resistance limits improving clock delay/skew over long distances

Benefits of UltraScale Clock Architecture



- 32 centrally located global clock buffers
- Root always in center of device
- Skew accumulates from center to edge
- 192 – 720 distributed global buffers
- Root can be in any clock region
- Balanced skew per clock network

Routing Challenge

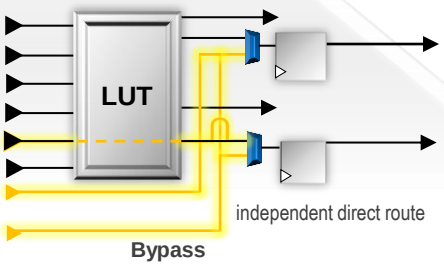
1. Routing demand
   - As busses widen the quantity and length of routing demand increases
2. Control Sets
   - More complex designs and clock gating cause design spreading with lower part utilization and longer route lengths
3. Metal Resistance
   - As metal layers shrink metal resistance limits improving routing delays over longer distances

CLB Innovations Enable Tighter Packing

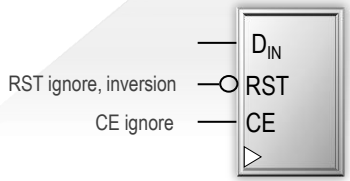
Removed slice boundary & added MUX
- Wider functions per block
- 2x distributed RAM density

Wider carry chain
- Wider functions per block

Dedicated inputs for each flip-flop
- Higher performance
- Reduces LUT utilization

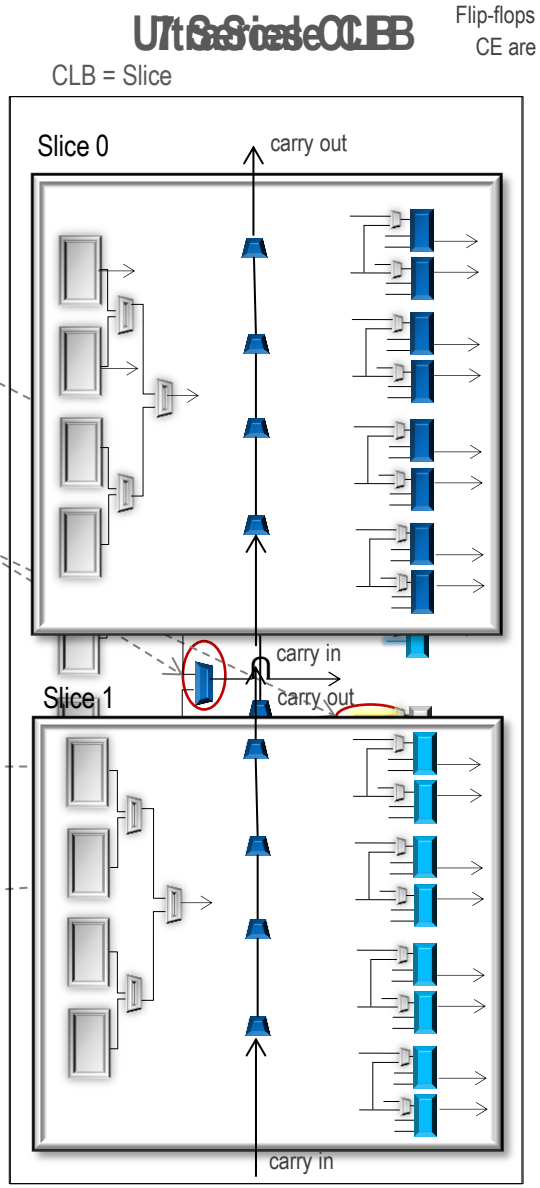


CE ignore and RST ignore, RST inversion
- Higher performance
- Eliminates synchronous reset bottlenecks



2x the number of CE's
- Improves CLB packing

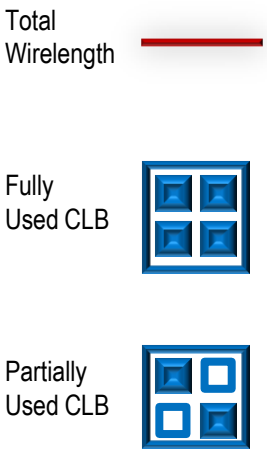
Flip-flops sharing same CE are same color

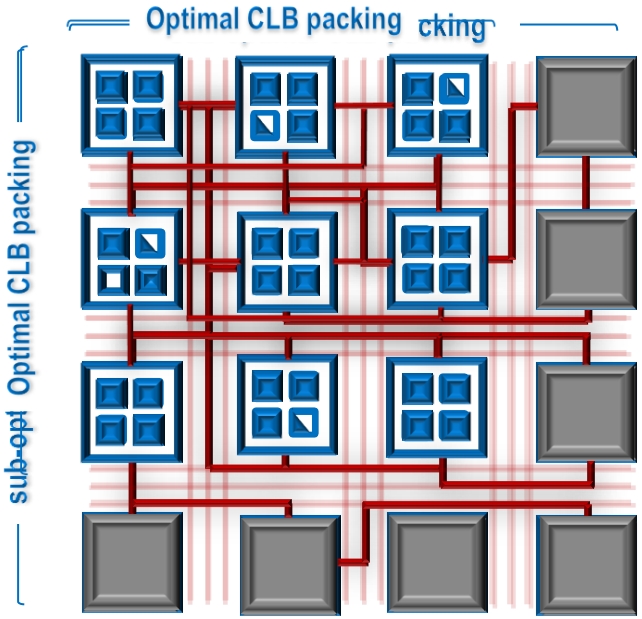


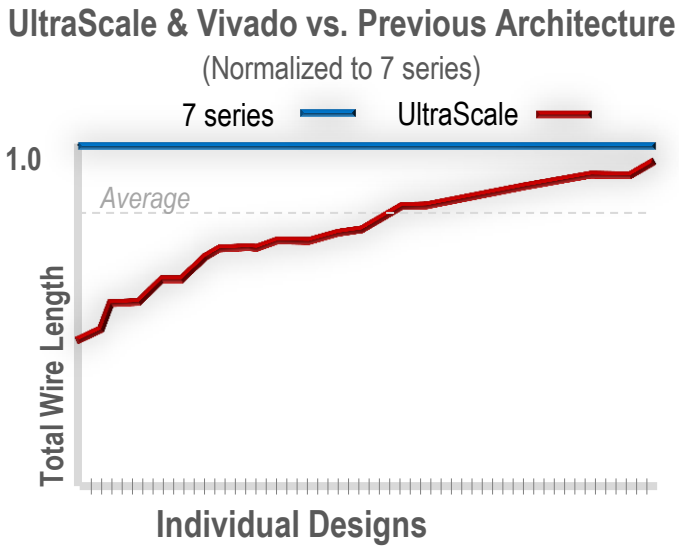
Flip-flops sharing same CE are same color

Design Use Fewer CLBs, Shorter Wire Length

- **Re-architected for flexible connectivity and efficient logic packing**
  - Better local routability, more control set flexibility, improved P&R algorithms
- **Tighter packing results in shorter net delays, less wire switching**
- **Translates into higher utilization, greater performance, less power**





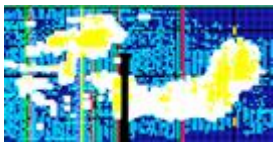


Additional Metal Layers

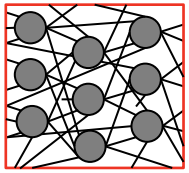
1. More and longer routing tracks
2. More thicker metal layers gives improved RC so that metal speed keeps up with transistor speed
3. Optimized wire lengths take best advantage of available metal

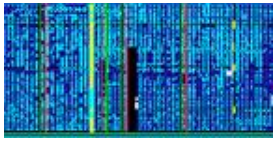
Routability Benchmark Suite

Metrics: Design Routing Complexity and Design Size

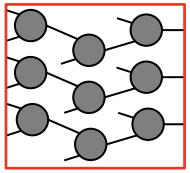


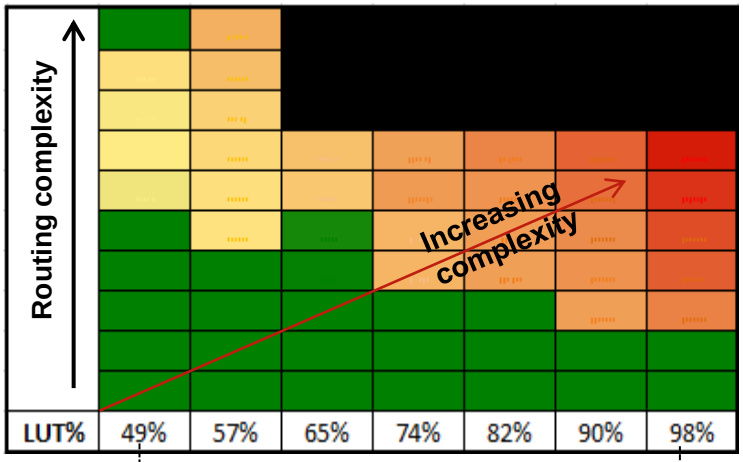
Congestion Metrics displayed in Vivado

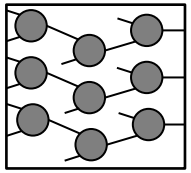


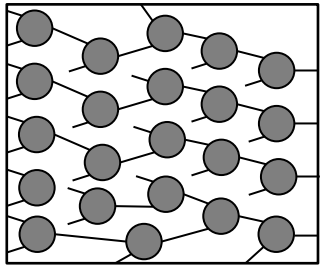


No congestion



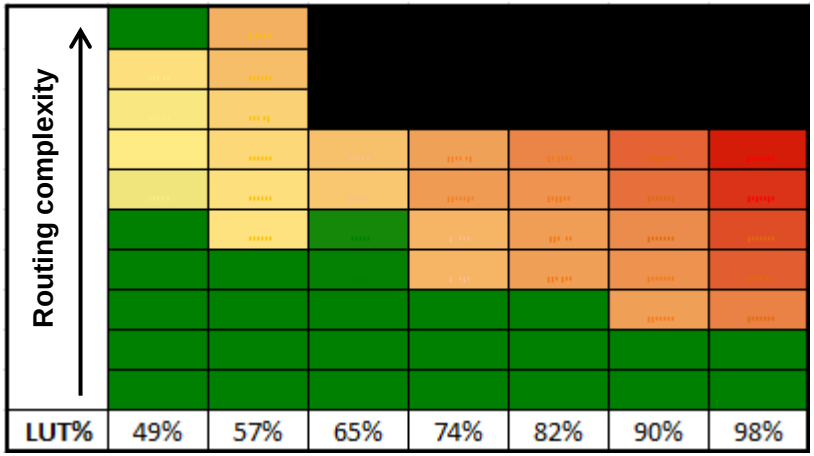




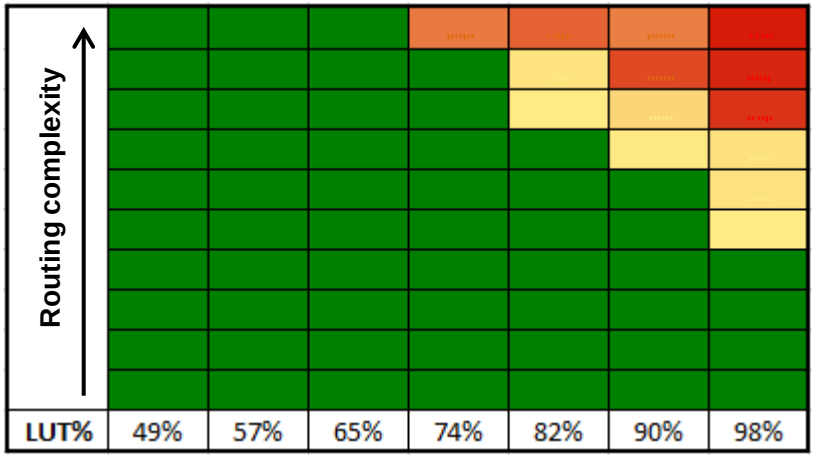


UltraScale Results

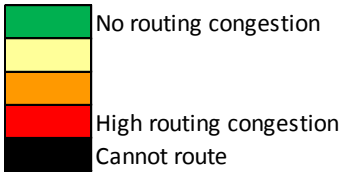
- Vivado® routes more complex designs on UltraScale
- UltraScale shows lower congestion on complex designs
- As a result, timing closure is accelerated
- Delivers 1 speedgrade higher Fmax



Classical FPGA Architecture

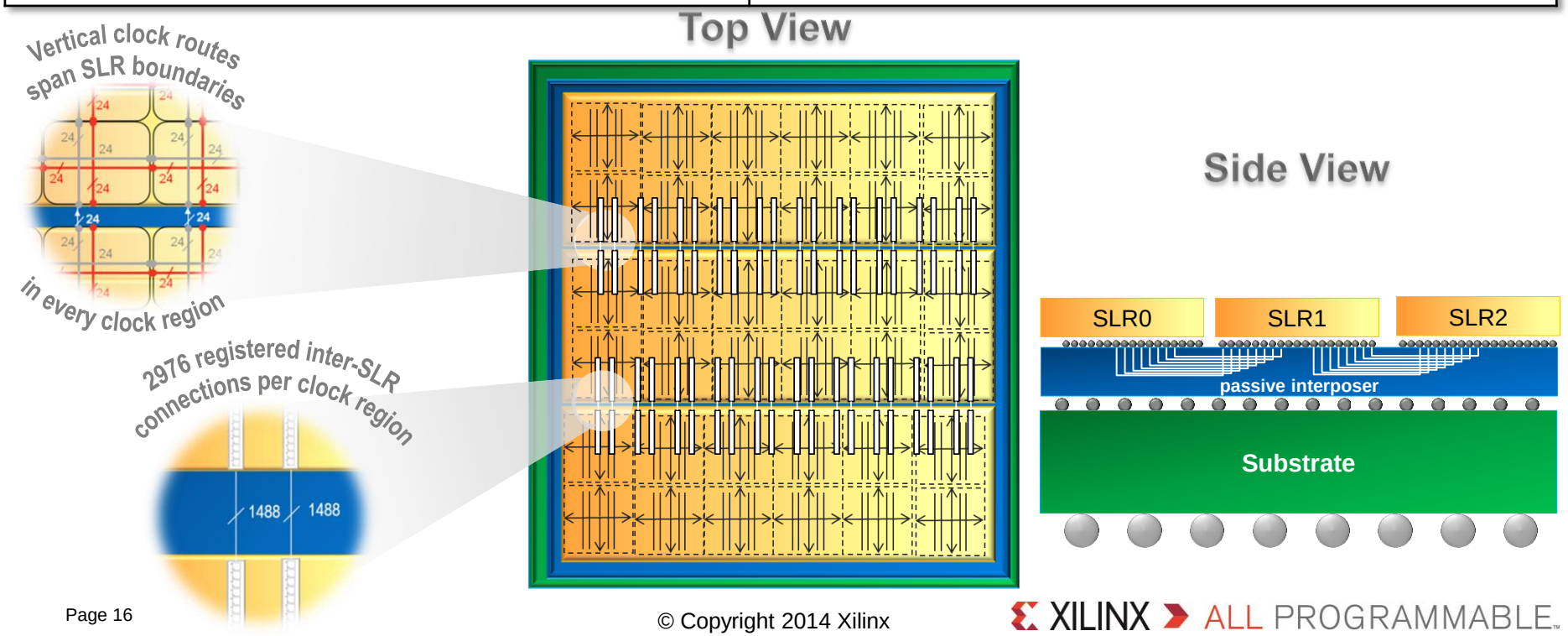


UltraScale

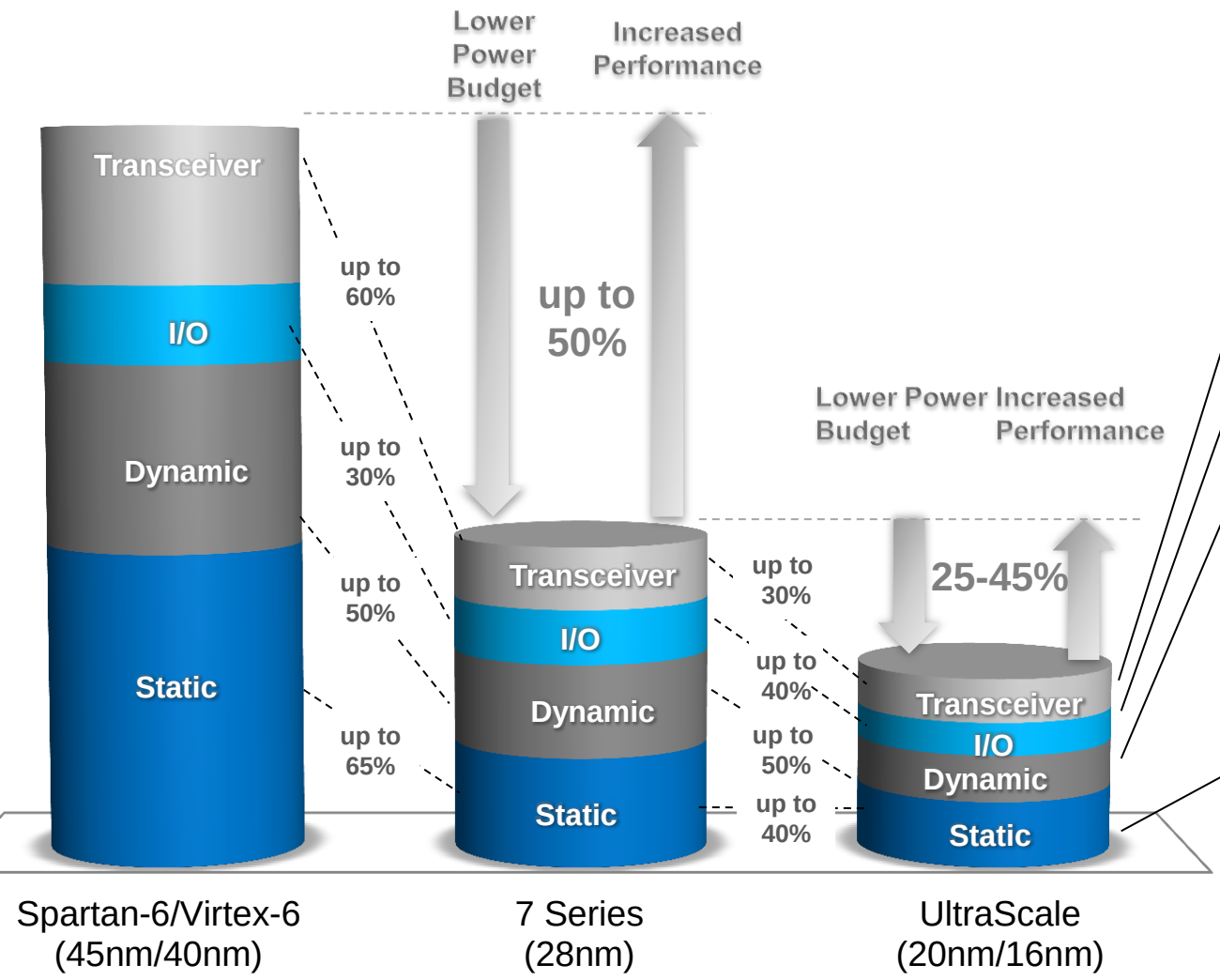


2nd Generation Stacked Silicon

Feature	Benefit
~20,000 registered routing lines between die	• Enables >500 MHz datapath performance between SLRs • Deterministic, predictable timing • Device size up to 4.4 million logic cells
Clocking Architecture Spans SLR boundaries	Abundant clock resources to meet demanding application
Foot-print compatibility between SSI and non-SSI devices	Ability to seamlessly migrate from monolithic to 3D-IC devices



Power Optimizations



UltraSCALE™ Architecture

- Architectural optimizations
- Low power mode

- I/O multi-mode control (cont'd from 28nm)
- DDR4 voltage reduction

- CLB packing & reduced wire length
- HW based clock gating on leaf cells
- BRAM hardened data cascading
- BRAM dynamic power gating
- DSP hardened features
- MMCM & PLL lower supply voltage

- Process node
- Power binning & lower voltage scaling
- 3D IC static power binned slices

Kintex® UltraScale™ FPGAs

	Part Number	XCKU035	XCKU040	XCKU060	XCKU075	XCKU100	XCKU115
Logic Resources	Logic Cells	355,474	424,200	580,440	756,000	958,440	1,160,880
	CLB Flip-Flops	406,256	484,800	663,360	864,000	1,095,360	1,326,720
	CLB LUTs	203,128	242,400	331,680	432,000	547,680	663,360
Memory Resources	Maximum Distributed RAM (Kb)	5,908	7,050	9,180	7,290	12,825	18,360
	Block RAM/FIFO w/ECC (36 Kb each)	540	600	1,080	1,188	1,680	2,160
	Block RAM/FIFO (18 Kb each)	1,080	1,200	2,160	2,376	3,360	4,320
	Total Block RAM (Mb)	19.0	21.1	38.0	41.8	59.1	75.9
Clock Resources	CMT (1 MMCM, 2 PLLs)	8	8	12	14	24	24
	I/O DLL	40	40	48	64	64	64
I/O Resources	Maximum Single-Ended HP I/Os	416	416	520	624	676	676
	Maximum Differential HP I/O Pairs	192	192	240	288	312	312
	Maximum Single-Ended HR I/Os	104	104	104	104	156	156
	Maximum Differential HR I/O Pairs	48	48	48	48	72	72
Integrated IP Resources	DSP Slices	1,700	1,920	2,760	2,592	4,200	5,520
	System Monitor	1	1	1	1	2	2
	PCIe® Gen1/2/3	2	3	3	4	6	6
	Interlaken	0	0	0	2	0	0
	100G Ethernet	0	0	0	1	0	0
	GTH 16 Gb/s Transceivers	16	20	32	52	64	64
Speed Grades	Commercial	-1	-1	-1	-1	-1	-1
	Extended	-2 -3	-2 -3	-2 -3	-2 -3	-2 -3	-2 -3
	Industrial	-1 -1L -2	-1 -1L -2	-1 -1L -2	-1 -1L -2	-1 -1L -2	-1 -1L -2

	Package Footprint	Package Dimensions (mm)	HR I/O, HP I/O, GTH 16 Gb/s					
Footprint Compatible with Virtex UltraScale	A676	27x27	104, 208, 16	104, 208, 16				
	A900	31x31	104, 364, 16	104, 364, 16				
	A1156	35x35	104, 416, 16	104, 416, 20	104, 416, 28	104, 416, 28		
	A1517	40x40			104, 520, 32	104, 520, 48	104, 520, 48	104, 520, 48
	B1517	40x40					104, 260, 64	104, 260, 64
	A1760	42.5x42.5				104, 624, 52		
	A1760						104, 624, 52	104, 624, 52
	D1924	45x45					156, 676, 52	156, 676, 52
	F1924	45x45					104, 624, 64	104, 624, 64

Virtex® UltraScale™ FPGAs

	Part Number	XCVU065	XCVU080	XCVU095	XCVU125	XCVU160	XCVU190	XCVU440
	Logic Cells	626,640	780,000	940,800	1,253,280	1,621,200	1,879,920	4,407,480
	CLB Flip-Flops	716,160	891,424	1,075,200	1,432,320	1,852,800	2,148,480	5,037,120
	CLB LUTs	358,080	445,712	537,600	716,160	926,400	1,074,240	2,518,560
Memory Resources	Maximum Distributed RAM (Kb)	4,230	3,980	4,800	8,460	10,710	12,690	28,710
	Block RAM/FIFO w/ECC (36 Kb each)	1,260	1,421	1,728	2,520	3,276	3,780	2,520
	Block RAM/FIFO (18 Kb each)	2,520	2,842	3,456	5,040	6,552	7,560	5,040
	Total Block RAM (Mb)	44.3	50.0	60.8	88.6	115.2	132.9	88.6
Clock Resources	CMT (1 MMCM, 2 PLLs)	10	16	16	20	26	30	30
	I/O DLL	40	64	64	80	104	120	120
	Fractional PLL	5	8	8	10	13	15	0
I/O Resources	Maximum Single-Ended HP I/Os	468	780	780	936	988	988	1,404
	Maximum Differential HP I/O Pairs	216	360	360	432	456	456	648
	Maximum Single-Ended HR I/Os	52	52	52	104	52	52	52
	Maximum Differential HR I/O Pairs	24	24	24	48	24	24	24
Integrated IP Resources	DSP Slices	600	672	768	1,200	1,560	1,800	2,880
	System Monitor	1	1	1	2	3	3	3
	PCIe® Gen1/2/3	2	4	4	4	4	6	6
	Interlaken	3	6	6	6	9	9	0
	100G Ethernet	3	4	4	6	7	9	3
	GTH 16 Gb/s Transceivers	20	32	32	40	52	60	48
	GTY 33 Gb/s Transceivers	20	32	32	40	52	60	0
Speed Grades	Commercial	-1	-1	-1	-1	-1	-1	-1
	Extended	-2 -3	-2 -3	-2 -3	-2 -3	-2 -3	-2 -3	-2 -3
	Industrial	-1 -1L -2	-1 -1L -2	-1 -1L -2	-1 -1L -2	-1 -1L -2	-1 -1L -2	-1 -1L -2

	Package Footprint [1]	Package Dimensions (mm)	HR I/O, HP I/O, GTH 16 Gb/s, GTY 33 Gb/s						
Footprint Compatible with Kintex® UltraScale Devices	C1517	40x40	52, 468, 20, 20	52, 468, 24, 24	52, 468, 24, 24				
	B1517	40x40		52, 312, 32, 32	52, 312, 32, 32	52, 312, 40, 32			
	A1760	42.5x42.5		52, 676, 32, 16	52, 676, 32, 16	52, 676, 36, 16			
	D1924	45x45		52, 780, 28, 24	52, 780, 28, 24	52, 780, 28, 24	-		
	E1924	45x45 47.5x47.5[2]		52, 624, 32, 32	52, 624, 32, 32	52, 624, 36, 36	52, 624, 36, 36	52, 624, 36, 36	
	J1924	45x45 47.5x47.5[2]			52, 312, 32, 32	52, 312, 40, 40	52, 312, 52, 52	52, 312, 52, 52	
	A2377	50x50				104, 936, 28, 24	52, 988, 28, 24	52, 988, 28, 24	
	B2377	50x50							52, 1248, 36, 0
	C2377	50x50						0, 520, 60, 60	
	A2892	55x55							52, 1404, 48, 0

Key Takeaways

- **Increased IO bandwidth**
  - Up to 120 transceivers
  - Up to 33 Gb/s per channel
- **Routing architecture and Vivado designed to reduce congestion and improve performance**
- **Clocking architecture for flexibility and performance**
- **2nd generation SSI delivers "more than Moore"**
  - Up to 4.4 million logic cells in largest device

Breakthrough architecture to address the most complex designs

Thank You!